Milk- and solid-feeding practices and daycare attendance are associated with differences in bacterial diversity, predominant communities, and metabolic and immune function of the infant gut microbiome

Amanda L. Thompson, *University of North Carolina at Chapel Hill*
Andrea Monteagudo-Mera, *University of North Carolina at Chapel Hill*
Maria B. Cadenas, *University of North Carolina at Chapel Hill*
Michelle Lampl, *Emory University*
M. A. Azcarate-Peril, *University of North Carolina at Chapel Hill*

**Journal Title:** Frontiers in Cellular and Infection Microbiology  
**Volume:** Volume 5  
**Publisher:** Frontiers Media | 2015-02-05  
**Type of Work:** Article | Final Publisher PDF  
**Publisher DOI:** 10.3389/fcimb.2015.00003  
**Permanent URL:** [https://pid.emory.edu/ark:/25593/rxs83](https://pid.emory.edu/ark:/25593/rxs83)

*Final published version: [http://dx.doi.org/10.3389/fcimb.2015.00003](http://dx.doi.org/10.3389/fcimb.2015.00003)*

**Copyright information:**  
© 2015 Thompson, Monteagudo-Mera, Cadenas, Lampl and Azcarate-Peril. This is an Open Access work distributed under the terms of the Creative Commons Attribution 4.0 International License ([https://creativecommons.org/licenses/by/4.0/](https://creativecommons.org/licenses/by/4.0/)).

*Accessed April 9, 2019 9:39 AM EDT*
Milk- and solid-feeding practices and daycare attendance are associated with differences in bacterial diversity, predominant communities, and metabolic and immune function of the infant gut microbiome

Amanda L. Thompson1, Andrea Monteagudo-Mera2, Maria B. Cadenas2, Michelle L. Lampl3 and M. A. Azcarate-Peril2,4*

1 Department of Anthropology, University of North Carolina, Chapel Hill, NC, USA
2 Microbiome Core Facility, Center for Gastrointestinal Biology and Disease, University of North Carolina, Chapel Hill, NC, USA
3 Department of Anthropology and Center for the Study of Human Health, Emory University, Atlanta, GA, USA
4 Department of Cell Biology and Physiology, School of Medicine, University of North Carolina, Chapel Hill, NC, USA

The development of the infant intestinal microbiome in response to dietary and other exposures may shape long-term metabolic and immune function. We examined differences in the community structure and function of the intestinal microbiome between four feeding groups, exclusively breastfed infants before introduction of solid foods (EBF), non-exclusively breastfed infants before introduction of solid foods (non-EBF), EBF infants after introduction of solid foods (EBF+S), and non-EBF infants after introduction of solid foods (non-EBF+S), and tested whether out-of-home daycare attendance was associated with differences in relative abundance of gut bacteria. Bacterial 16S rRNA amplicon sequencing was performed on 49 stool samples collected longitudinally from a cohort of 9 infants (5 male, 4 female). PICRUSt metabolic inference analysis was used to identify metabolic impacts of feeding practices on the infant gut microbiome. Sequencing data identified significant differences across groups defined by feeding and daycare attendance. Non-EBF and daycare-attending infants had higher diversity and species richness than EBF and non-daycare attending infants. The gut microbiome of EBF infants showed increased proportions of *Bifidobacterium* and lower abundance of *Bacteroidetes* and *Clostridiales* than non-EBF infants. PICRUSt analysis indicated that introduction of solid foods had a marginal impact on the microbiome of EBF infants (24 enzymes overrepresented in EBF+S infants). In contrast, over 200 bacterial gene categories were overrepresented in non-EBF+S compared to non-EBF infants including several bacterial methyl-accepting chemotaxis proteins (MCP) involved in signal transduction. The identified differences between EBF and non-EBF infants suggest that breast milk may provide the gut microbiome with a greater plasticity (despite having a lower phylogenetic diversity) that eases the transition into solid foods.

**Keywords:** infant gut microbiome, breastfeeding, metagenomics, daycare, feeding transitions

**INTRODUCTION**

During the first year of life the human infant gut microbiome undergoes rapid maturation, changing dynamically in response to early environmental exposures, such as delivery type, hygiene measures, and diet (Guarner and Malagelada, 2003). Infant feeding practices, particularly breast- and formula-feeding, have been shown to influence the structure and function of this developing microbiome. Human milk provides the infant with a rich microbial consortium and a variety of oligosaccharides, prebiotics ensuring gut colonization by microbes beneficial for metabolism and immune development (Jeurink et al., 2013). Differential development of the intestinal microbiome in response to human milk may underlie documented differences in infectious illness morbidity, allergy, and obesity risk between breast- and formula-fed infants (Fanaro et al., 2003; Ip et al., 2007; Donnet-Hughes et al., 2010).

The gut microbiota of breastfed infants has been characterized as having less diverse colonization and greater proportions of *Bifidobacterium* and *Lactobacillus* species compared to non-breastfed infants. A recent study by Tannock et al. (2013) identified Bifidobacteriaceae (61%), Enterobacteriaceae (8%), and Coriobacteriaceae (6%) as the most abundant bacterial families detected in stool from breastfed babies. The microbiome of formula-fed infants is more diverse with greater proportions of *Bacteroides*, *Clostridium* and enterobacteria (Orrhage and Nord, 1999; Vael and Desager, 2009; Marques et al., 2010). The presence
and/or predominance of *Bifidobacterium* in exclusively breast-fed compared to non-exclusively breastfed or formula-fed infants is still under debate. While numerous studies have reported a lower abundance of *Bifidobacterium* in formula-fed infants relative to breast-fed infants (Yoshioka et al., 1983; Favier et al., 2002; Hopkins et al., 2005), other studies have reported *Bifidobacterium* as the dominant taxa in both groups (Gomez-Llorente et al., 2013; Tannock et al., 2013; Yatsunenko et al., 2012). The predominant families in mixed-fed infants have been less well characterized after the neonatal period, but their microbiomes have been assumed to be intermediate between those of breast- and formula-fed infants (Orrhage and Nord, 1999).

The composition of the early gut microbiome changes in both breast- and formula-fed infants as solids foods are introduced into the diet (Gunner and Malagelada, 2003; Amarri et al., 2006; Nielsen et al., 2007; Palmer et al., 2007). The transition from a milk-based to a solids-based diet exposes infants to novel non-digestible plant carbohydrates, animal protein, and fats providing new substrates for the survival and dominance of bacterial species not supported by breastmilk and/or formula (Parrett and Edwards, 1997; Fallani et al., 2011). Introduction of solid foods has been associated with increased populations of *Bacteroides* (Koenig et al., 2011) and decreased populations of bifidobacteria, enterobacteria, and some *Clostridium* species (Fallani et al., 2011). In older children and adults, dietary composition, particularly the balance between carbohydrates and protein/animal fat, is associated with microbiome changes with greater *Bacteroides* abundance associated with diets higher in protein and animal fat and *Prevotella* abundance associated with greater carbohydrate intake relative to meat and dairy (De Filippo et al., 2010; Wu et al., 2011). Along with the increased abundance of *Bacteroides*, changes in bacterial populations caused by introduction of solid foods have been correlated with increased short chain fatty acids (SCFAs) levels (Koenig et al., 2011). Such diet-induced differences in the gut microbiota may contribute to host metabolism and immune function by regulating genes involved in lipid and carbohydrate metabolism, altering endocrine functions, increasing inflammatory responses, and influencing energy balance and body weight (Gill et al., 2006; Greiner and Backhed, 2011; Nauta et al., 2013). Yet, few studies have examined the effects of the solid feeding transition on the composition of the gut microbiota or its function longitudinally in a cohort of healthy, predominantly breastfed infants with well-documented diet, growth, and morbidity information.

To investigate the impact of infant feeding patterns (exclusive breastfeeding and mixed-feeding of breastmilk and formula) and the subsequent introduction of solid foods on the development and function of the gut microbiome, we analyzed the microbiome of stool samples (*N* = 49) collected longitudinally from 9 American infants, ages 2 weeks to 14 months, by 16S rRNA amplicon pyrosequencing. We described the species diversity associated with exclusive breastfeeding (EBF) and mixed breast and formula feeding (non-exclusively breast-fed, non-EBF) and tested whether the introduction of solid foods was associated with differences in bacterial composition in these groups. Additionally, we used PICRUSt (Langille et al., 2013) to construct a community-level metabolic network of the microbiome and compared the abundance of pathways across feeding groups to identify metabolic differences associated with feeding patterns and transitions.

As a secondary aim, we tested for significant differences in the gut microbiome by out-of-home daycare attendance. Young children in out-of-home daycare settings have higher risk of gastrointestinal and respiratory infections (Augustine et al., 2013). A number of recent trials have shown that probiotic supplementation reduces the incidence and duration of diarrheal episodes in children in out-of-home care, suggesting that this morbidity may be due to differences in the composition and function of the microbiome (Gutierrez-Castrellon et al., 2014). Thus, we examined whether daycare attendance was associated with differences in the microbiome beyond those attributable to feeding practices.

**MATERIALS AND METHODS**

**SAMPLE**

Stool samples (*n* = 49) were collected from 9 infants (5 males, 4 females) participating in a prospective, mixed-longitudinal study investigating infant growth and development during the first year of life (Thompson and Lampl, 2013). As previously described, infants in the parent study were recruited from daycare centers, lactation support groups, and Atlanta-area universities. At recruitment, infants were aged 7 days to 10.5 months (median age at entry 19.6 weeks, IQR = 10.8–32.8 weeks) and were followed weekly for a median of 16 weeks (IQR = 16–20.8 weeks). All participating infants were born after 37 gestational weeks with birthweights >2500 g. Infants were chosen for the current analytic sample if they had a minimum of two retained stool samples with at least one preceding and one following a feeding transition, such as supplementation with formula or introduction of solid foods. Stool samples in the current study were from infants aged 13 days to 14 months of age.

Characteristics of infants in the analytic sample are described in Table 1. Eight of nine infants received breastmilk, four of nine infants were exclusively breastfed during at least one of the sample collection periods, and one infant received formula and solids with no breastmilk during the study period. Infants were breastfed for a median of 9 months (range: 5.2–13 months). Mixed fed infants were all predominantly breastfed, receiving the equivalent of one to two bottles of formula per day. Solid foods were introduced at a median age of 5 months (range: 4–6 months). Four of nine infants attended out-of-home daycare. Two of the nine infants in daycare and none of the five infants cared for in-home received a narrow spectrum oral antibiotic (Cefdinir) for ear infections in the week preceding sample collection. Infants in the analytic sample did not differ significantly from infants in the full sample in age, gender, or feeding type. Recruitment and data collection protocols were approved by the Emory University Institutional Review Board. Secondary analysis was approved by the University of North Carolina at Chapel Hill Institutional Review Board.

**Stool sample collection**

Parents were asked to retain soiled diapers after each diaper change during a 24-h period. The diapers were placed in individual plastic storage bags labeled for date and time of collection.
Table 1 | Characteristics of participating infants at the time of stool sample collection.

<table>
<thead>
<tr>
<th>Infant</th>
<th>Sex</th>
<th>Age (days)</th>
<th>Age (months)</th>
<th>Delivery type</th>
<th>Daycare</th>
<th>Any BMb</th>
<th>Any Formulae</th>
<th>EBFd</th>
<th>Any Solids e</th>
<th>Age at Solid Introduction (months)</th>
<th>Antibiotic Usef</th>
<th>Birthweight (kg)</th>
<th>BMI (kg/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>F</td>
<td>25</td>
<td>1</td>
<td>C-section</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>5</td>
<td>No</td>
<td>3.09</td>
<td>14.0</td>
</tr>
<tr>
<td>73</td>
<td>2</td>
<td>118</td>
<td>4</td>
<td></td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>14.8</td>
<td></td>
</tr>
<tr>
<td>156</td>
<td>5</td>
<td>177</td>
<td>6</td>
<td></td>
<td>No</td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>4</td>
<td>Yes</td>
<td>16.1</td>
<td></td>
</tr>
<tr>
<td>210</td>
<td>7</td>
<td>289</td>
<td>10</td>
<td></td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>15.1</td>
<td></td>
</tr>
<tr>
<td>396</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>14.7</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>13</td>
<td>0</td>
<td>Vaginal</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>4</td>
<td>No</td>
<td>2.72</td>
<td>9.7</td>
</tr>
<tr>
<td>53</td>
<td>2</td>
<td>84</td>
<td>3</td>
<td></td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12.7</td>
<td></td>
</tr>
<tr>
<td>181</td>
<td>6</td>
<td>244</td>
<td>8</td>
<td></td>
<td>No</td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>4</td>
<td>No</td>
<td>14.1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>110</td>
<td>4</td>
<td>C-section</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>6</td>
<td>No</td>
<td>3.96</td>
<td>15.6</td>
</tr>
<tr>
<td>124</td>
<td>4</td>
<td>155</td>
<td>5</td>
<td></td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>12.0</td>
<td></td>
</tr>
<tr>
<td>167</td>
<td>5</td>
<td>189</td>
<td>6</td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>13.0</td>
<td></td>
</tr>
<tr>
<td>252</td>
<td>8</td>
<td>305</td>
<td>10</td>
<td></td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>13.6</td>
<td></td>
</tr>
<tr>
<td>357</td>
<td>12</td>
<td>417</td>
<td>14</td>
<td></td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>19.5</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>M</td>
<td>80</td>
<td>3</td>
<td>Vaginal</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>6</td>
<td>No</td>
<td>3.83</td>
<td>15.8</td>
</tr>
<tr>
<td>110</td>
<td>4</td>
<td>141</td>
<td>5</td>
<td></td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>15.7</td>
<td></td>
</tr>
<tr>
<td>191</td>
<td>6</td>
<td>196</td>
<td>6</td>
<td></td>
<td>No</td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>6</td>
<td>No</td>
<td>15.3</td>
<td></td>
</tr>
<tr>
<td>220</td>
<td>7</td>
<td>253</td>
<td>8</td>
<td></td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>15.7</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>F</td>
<td>103</td>
<td>3</td>
<td>Vaginal</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>5</td>
<td>No</td>
<td>3.66</td>
<td>14.5</td>
</tr>
<tr>
<td>217</td>
<td>7</td>
<td>252</td>
<td>8</td>
<td></td>
<td>No</td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>5</td>
<td>No</td>
<td>16.2</td>
<td></td>
</tr>
<tr>
<td>279</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>17.2</td>
<td></td>
</tr>
<tr>
<td>103</td>
<td>M</td>
<td>213</td>
<td>7</td>
<td>Vaginal</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>5.5</td>
<td>No</td>
<td>3.69</td>
<td>15.9</td>
</tr>
<tr>
<td>255</td>
<td>8</td>
<td>283</td>
<td>9</td>
<td></td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
<td>No</td>
<td>15.3</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
Bagged diapers were stored in portable coolers equipped with ice packs frozen at −80°C. Diapers were transported to the Laboratory of Reproductive Ecology at Emory University for storage at −80°C within 24-h of collection. Retained portions of the original samples were analyzed in the current study.

Infant feeding
Infant feeding type was classified on a weekly-basis using 24-h diet histories where parents were asked to record all the foods and liquids, including specific types and amounts, consumed in the 24-h preceding sample collection. For the present analysis, infants were classified as exclusively breastfed (EBF) if they only received breastmilk and small amounts of medicine or vitamins. Since no infants received formula exclusively, infants were classified as non-exclusively breastfed (non-EBF) if they received both breastmilk and small amounts of medicine or vitamins. Data were categorized into two groups based on whether or not they received infant formula in addition to solids: (1) breastmilk with solids (EBF+S) and (2) formula with solids with or without breastmilk (non-EBF+S).

DNA isolation
Isolation of total DNA was performed using a customized protocol in the Qiagen BioRobot Universal (Qiagen) with the Qiagen Blood and Tissue and QIAmp DNA Stool protocols. The Qiagen protocol was modified to ensure isolation of DNA from Gram positive as well as Gram negative bacteria as follows: approximately 100 mg of feces were resuspended in 1.4 ml of ASL buffer, homogenized in a TissueLyser (Qiagen) for 2 min at 25 Hz, and heated for 5 min at 95°C. Samples were then vortexed (15 s) and centrifuged, and the supernatants were transferred to a new tube. An InhibitEX tablet (Qiagen) was added to remove inhibitors, the mixtures were incubated for 1 h at room temperature, and centrifuged (20,000 × g, 3 min). Two hundred µl of the supernatant were then transferred to a new Eppendorf tube containing 15 µl of proteinase K (600 mAU/ml) and 200 µl of buffer AL were added to the mixture, which was incubated at 70°C for 10 min. Two hundred µl of ethanol were then added and the mix was transferred to the bed of the BioRobot Universal and DNA purification was carried out using a customized isolation protocol. DNA was visualized by electrophoresis and quantified using Quanti-IT™ PicoGreen® dsDNA Reagent (Molecular Probes, Life Technologies division).

16S rDNA bacterial amplicon pyrosequencing
Initial amplification of the V1-V2 region of the bacterial 16S gene was performed on 50 individual samples from the study as previously described (Devine et al., 2013). Briefly, master mixes for amplicon generation contained the Qiagen Hotstar Hi-Fidelity Polymerase Kit (Qiagen, Valencia CA) with a forward primer composed of the Roche Titanium Fusion Primer A (5′-CCATCTCATCCCTGCGTGTCTCAG-3′), a 10 bp Multiplex Identifier (MID) sequence (Roche, Indianapolis, IN) unique to each of the samples and the universal bacteria primer 8F (5′-AGAGTTTGATCTCGTGGTCAAG-3′) (Edwards et al., 1989). The reverse primer was composed of the Roche Titanium Primer B (5′-CCTATCCCCCTGTGCGCTGAGTTC-3′), the identical 10 bp MID sequence as the forward primer and the reverse bacteria primer 338R (5′-GCTGCGCCGCTGCTAG-3′) (Fierer et al., 2008). In order to ensure detection of Bifidobacterium species normally highly represented in breastfed babies we used a combination (4:1) of the primers 8F and Bifidobacterium-specific Bifido-8F (5′-AGGTTCGATTCTGGCTCAG-3′) (Martinez et al., 2010) with 335R as the reverse primer as described (Davis et al., 2011). The thermal profile for the amplification of each sample was an initial denaturing step at 94°C for 5 min, followed by a cycling of denaturation at 94°C for 45 s, annealing at 50°C for 30 s, a 1 min 30 s extension at 72°C (35 cycles), a 10 min extension at 72°C and a final hold at 4°C. Each sample was gel-purified individually using the E-Gel Electrophoresis System (Invitrogen, Life Technologies division), quantified using Quanti-IT™ PicoGreen® dsDNA Reagent, and the concentration was standardized. The 16S rDNA amplicons from the pooled sample were sequenced on a Roche GS FLX Titanium instrument (Microbiome Core Facility, Chapel Hill NC). Initial data analysis, base pair calling

### Table 1 | Continued

<table>
<thead>
<tr>
<th>Infant</th>
<th>Sex</th>
<th>Age (days)</th>
<th>Age (months)</th>
<th>Delivery type</th>
<th>Any BM</th>
<th>Any Formula</th>
<th>EBF</th>
<th>Any Solids</th>
<th>Age at Solid Introduction (months)</th>
<th>Antibiotic Use</th>
<th>Birthweight (kg)</th>
<th>BMI (kg/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>111</td>
<td>M</td>
<td>340</td>
<td>11</td>
<td>Vaginal</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>4.5</td>
<td>No</td>
<td>3.01</td>
<td>14.9</td>
</tr>
<tr>
<td>114</td>
<td>M</td>
<td>323</td>
<td>11</td>
<td>Vaginal</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>4.5</td>
<td>No</td>
<td>3.77</td>
<td>15.7</td>
</tr>
</tbody>
</table>

*a Number indicates sample id in original study.
*b Indicates whether infant received any breastmilk during the week prior to sample collection.
*c Indicates whether infant received breastmilk during the week prior to sample collection.
*d Indicates whether infant was exclusively breastfed (receiving only breastmilk and small quantities of medications) during the week prior to sample collection.
*e Indicates whether infant received antibiotics in the week prior to sample collection.
*f Indicates whether infant received solid foods during the week prior to sample collection.
*g Indicates whether infant received solid foods in the week prior to sample collection. No mothers reported antibiotic use during the study period.
and trimming to yield high quality reads, were performed by Research Computing at UNC.

**Amplicon high-throughput sequencing data analysis**

This bioproject has been registered at the National Center for Biotechnology Information (NCBI) under BioProject ID PRJNA270898. Analysis of sequencing data was carried out using the QIIME pipeline (Caporaso et al., 2010) as described (Devine et al., 2013). The combined raw sequencing data plus metadata describing the samples were de-multiplexed and filtered for quality control. Next, data was denoised using the recommended denoising method (Denoiser, Caporaso et al., 2010). Sequences were grouped into OTUs (Operational Taxonomic Units) at a 97% level using Uclust (Edgar, 2010). After taxonomic assignation of OTUs, sequences were aligned and phylogenetic trees were built (Price et al., 2010). QIIME was also used to calculate alpha diversity on rarefied OTU tables to assess sampling depth coverage using observed species, Shannon, and phylogenetic diversity (PD) metrics. These metrics were used to quantify microbiome community structure with species richness (S) providing a measure of the number of different bacterial species present per sample and phylogenetic diversity (PD) providing a measure of the diversity of taxa present. To correct for different numbers of sequences in each sample, we randomly subsampled 2908 sequences in each sample (the number of sequences in the sample with the lowest number of sequences). To evaluate the similarities between bacterial communities, a combination of UniFrac significance and principal coordinate analysis (PCoA) using Fast UniFrac (Lozupone et al., 2006) were performed to compare samples based on relevant parameters (Table 2). Supervised machine-learning technique Random Forests (RF) within QIIME (Knights et al., 2011) were used to identify bacterial genera and species-level OTUs that differentiate the community composition between feeding groups.

**PICRUSt analysis of 16S amplicon sequencing data**

PICRUSt is a tool designed to infer metagenomic information from 16S amplicon sequencing data (Langille et al., 2013). In the present study, analysis of 16S amplicon sequencing data was performed using the default settings of PICRUSt (version 0.9.1). The resulting metagenomic data were entered into the HMP unified metabolic analysis network (HUMAnN) (Abubucker et al., 2012) pipeline (version 0.98) to sort individual genes into Kyoto encyclopedia of genes and genomes (KEGG) pathways representing varying proportions of each imputed sample metagenome.

**EXPERIMENTAL DESIGN**

Forty-nine stool samples from 9 infants were divided in four groups to investigate differences between EBF and non-EBF infants before and after the introduction of solid foods: (a) exclusively breastfed infants before introduction of solid foods (EBF, \( n = 12 \), infant age 25–191 days), (b) non-exclusively breastfed infants before introduction of solid foods (non-EBF, \( n = 6 \), infant age 13–155 days), (c) EBF infants after introduction of solid foods (EBF+S, \( n = 17 \), infant age 177–396 days), and (d) non-EBF infants after introduction of solid foods (non-EBF+S, \( n = 14 \), infant age 167–415 days). Secondarily, we tested whether, within these groups, infants differed by daycare attendance (in home vs. out of home care).

**Statistical analyses**

We used the non-parametric test ANOSIM (Analysis of Similarities) within QIIME to test whether two or more groups of samples were significantly different (Fierer et al., 2010). The number of permutations was set at 999 to calculate \( p \)-values. To identify bacterial taxa and biological pathways predicted by PICRUSt that were significantly (\( p \leq 0.05 \)) over or under represented in the different groups we applied the Steel-Dwass All Pairs test in JMP genomics (SAS JMP Genomics 5.0).

**RESULTS**

Amplicon pyrosequencing of stool samples yielded a total of 304,175 reads with an average of 6083 reads per sample. Sequences were assigned to 814 Operational Taxonomic Units (OTUs) at \( \geq 97\% \) similarity clustering into 110 genera, 20 classes, and 9 phyla. Bacterial relative abundance at phylum level, phylogenetic diversity (PD), bacterial gene functional representation, feeding type and collection time point for individual infants are shown in Supplemental Figure 2.

---

### Table 2 | Analysis of similarities (ANOSIM) between parameters and categories evaluated in this study.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Categories</th>
<th>( p )-value</th>
<th>( R )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Under 5, over 5 months</td>
<td>0.003</td>
<td>0.1608</td>
</tr>
<tr>
<td>Feeding1</td>
<td>Any breast milk (Yes-No)</td>
<td>0.180</td>
<td>0.1123</td>
</tr>
<tr>
<td>Feeding2</td>
<td>Any formula (Yes-No)</td>
<td>0.003</td>
<td>0.1579</td>
</tr>
<tr>
<td>Feeding3</td>
<td>Any solids (Yes-No)</td>
<td>0.002</td>
<td>0.1693</td>
</tr>
<tr>
<td>Feeding4</td>
<td>Exclusively breastfed (EBF) plus solids (Yes-No)</td>
<td>0.868</td>
<td>-0.0613</td>
</tr>
<tr>
<td>Feeding5</td>
<td>Non-exclusively breastfed (non-EBF) plus solids (Yes-No)</td>
<td>0.009</td>
<td>0.1784</td>
</tr>
<tr>
<td>Feeding Group</td>
<td>EBF, non-EBF, EBF+S, non-EBF+S</td>
<td>0.001</td>
<td>0.2394</td>
</tr>
<tr>
<td>Feeding by Age1</td>
<td>Exclusively breastfed younger than 5.5 months (Yes-No)</td>
<td>0.023</td>
<td>0.1846</td>
</tr>
<tr>
<td>Feeding by Age2</td>
<td>Non-exclusively breastfed younger than 5.5 months (Yes-No)</td>
<td>0.193</td>
<td>0.1041</td>
</tr>
<tr>
<td>Daycare</td>
<td>Yes-No</td>
<td>0.008</td>
<td>0.1870</td>
</tr>
<tr>
<td>Sex</td>
<td>Female, Male</td>
<td>0.011</td>
<td>0.0791</td>
</tr>
</tbody>
</table>

\( \alpha \) was set at 0.05.

Bold values indicate factors that had a weak (ANOSIM \( R < 0.2, p < 0.05 \)) or medium (ANOSIM \( R = 0.2 \) to 0.5, \( p < 0.05 \)) statistically significant impact on variables.
**BACTERIAL DIVERSITY WAS INFLUENCED BY AGE, FEEDING, AND DAYCARE**

Analysis of similarities (ANOSIM) showed statistically significant ($p < 0.05$) differences in the infant gut microbiome due to age, feeding, and daycare attendance (Table 2), indicating that the compositional dissimilarities between the groups were higher than those within the groups. Regardless of the feeding type, analysis of the stool microbiome composition of samples collected over time reflected an increment of the average bacterial Phylogenetic Diversity (PD) and species richness ($S$) during the first year. Infants aged 12–14 months had the highest number of species ($S = 71.4 \pm 29.5$, Supplemental Figure 1A) and diversity values ($PD = 8.76 \pm 2.5$, Supplemental Figure 1B).

We observed significant differences in species diversity and richness between the groups ($p < 0.05$ for all comparisons). EBF infants had the lowest species richness and diversity ($PD = 5.8 \pm 1.3$, $S = 35.3 \pm 11.2$). Species richness and diversity was higher in non-EBF infants ($PD = 6.4 \pm 1.8$, $S = 41.5 \pm 23.8$) and EBF+S infants ($PD = 6.9 \pm 1.1$, $S = 47.11 \pm 14.9$) reaching the highest values in non-EBF+S infants ($PD = 8.3 \pm 2.0$, $S = 71.9 \pm 18.7$) (Figures 1A,B).

Principal coordinates analysis (PCoA) of unweighted UniFrac (Lozupone and Knight, 2005) also identified significant differences in the infant stool microbiome according to feeding pattern. Samples from EBF infants segregated from samples from non-EBF infants (Figure 1C). This segregation continued after solid foods were introduced into the diet (Figure 1D). Although EBF samples significantly clustered away from EBF+S samples (Figure 1E), these clusters were less defined ($R = 0.18$) than those observed between non-EBF vs. non-EBF+S samples ($R = 0.22$; Figure 1F).

Infants in out-of-home daycare had significantly higher diversity and species richness values ($PD = 7.7 \pm 2.0$, $S = 63.1 \pm 20.4$) compared to infants cared for in home ($PD = 6.6 \pm 1.5$, $S = 44.8 \pm 19.2$, $p < 0.05$), regardless of the feeding type (Figures 2A,B). A significant clustering pattern was also observed in PCoA plots by daycare attendance (Figure 2C).

**DIFFERENCES IN THE COMPOSITION OF THE STOOL MICROBIOME OF EBF AND NON-EBF INFANTS**

The stool microbiome of EBF and non-EBF infants was characterized by different phyla proportions. Actinobacteria were significantly over represented in EBF infants (average relative abundance 48.3%, $p = 0.055$) while Bacteroidetes were over represented in non-EBF infants (42.6%, $p = 0.054$). No significant differences were detected between groups in the Firmicutes phylum (37.6% in EBF, 33% in non-EBF infants) (Figure 3A, Table 3). Within the phyla, several taxa differed significantly between feeding groups. Within the phylum Actinobacteria, the relative abundance of *Bifidobacterium* was lower in non-EBF infants than in EBF infants, while *Eggerthella* was more abundant in the non-EBF group. *Bacteroides* within the phylum Bacteroidetes was significantly more abundant in non-EBF infants, though these differences did not reach statistical significance. Within the phylum Firmicutes, abundance of Clostridiales, Lachnospiraceae, *Blautia* and *Faecalibacterium* was significantly higher ($p < 0.05$) in non-EBF compared to EBF infants (Table 3).

Random Forests (RF) analysis revealed distinct community signatures for EBF and non-EBF infants (baseline error = 0.24, cross-validation error = 0.21 ± 0.09). We considered an OTU to be highly predictive if its importance score was at least 0.001, and in our study 60 OTUs were identified as highly predictive (Supplemental Table 1). Of the 60 highly predictive OTUs, 17 corresponded to the phylum Actinobacteria and 12 specifically to the genus *Bifidobacterium*. Other highly predictive genera included *Streptococcus* (5 OTUs), *Veillonella* (3 OTUs), *Parabacteroides* (3 OTUs), *Eggerthella*, *Bacteroides*, *Actinomyces*, *Enterococcus*, and *Coprobacillus* (each with 2 OTUs).

**GUT MICROBIOME COMPOSITIONAL IMPACT OF THE INTRODUCTION OF SOLID FOODS**

The introduction of solid foods into the diet of EBF infants resulted in a marked, significant increase in the abundance of Bacteroidetes, especially *Bacteroides* with a concomitant, though non-statistically significant, reduction in the relative abundance of Firmicutes and Actinobacteria (Figure 3, Table 3). At the genus level we observed significantly decreased abundance of *Staphylococcus* and *Roseateles* and increased abundance of *Eggerthella*, *Blautia*, and *Peptostreptococcaceae* after incorporation of solid foods in EBF infants (Table 3). The introduction of solid foods into the diet of non-EBF infants was associated with an increased abundance of *Actinobacteria* ($p = 0.06$), specifically *Bifidobacterium* (Figure 3, Table 3). Within the phylum Firmicutes, *Ruminococcaceae* and *Blautia* increased significantly after the introduction of solids.

RF analysis identified distinct community signatures for EBF+S and non-EBF+S infants (baseline error = 0.45, cross-validation error = 0.12 ± 0.15). In both groups Actinobacteria was the dominant phylum, followed by Firmicutes and Bacteroides (Figure 3). However, analysis at the genus level showed that abundance of *Lactobacillus* and *Ruminococcaceae* was significantly higher in EBF+S infants. Non-EBF+S infants showed significantly higher levels of Clostridiales, *Blautia*, *Faecalibacterium*, *Anaerotruncus*, and *Eubacterium* (Figure 3, Table 3).

**A COHORT CORE MICROBIOME DEFINED BY BREASTFEEDING**

The core microbiome, defined as “organisms common across microbiomes hypothesized to play a key role in ecosystem function within a habitat” (Turnbaugh et al., 2007), contains members common to two or more microbial assemblages associated with a condition such as feeding pattern. In this study, the core microbiome in EBF infants included *Bifidobacterium* and *Coprobacillus*, which were present in 75% of the samples. In contrast, *Veillonella* and Clostridiales were present in 100% and *Bacteroides* in 75% of non-EBF infants (Figure 5). Incorporation of solid foods into the diet of EBF infants was translated into the presence, in 75% of samples, of *Veillonella*, *Roseburia* and the family Lachnospiraceae in addition to the persistent species identified in EBF samples. Introduction of solid foods to the diet of non-EBF infants resulted in the identification of *Bifidobacterium*, *Streptococcus*, and *Coprobacillus* in the microbiome of 75% of non-EBF infants.
Thompson et al. Infant feeding and the microbiome

FIGURE 1 | (A) A comparison of species richness S and (B) phylogenetic diversity PD between feeding groups. Non-EBF infants showed significantly higher S and PD values than EBF infants, and both values were impacted by the introduction of solid foods (*p < 0.05). (C–F) Comparison of unweighted UniFrac PCoA plots with repeated resampling (jackknifing) of microbiota from different feeding groups. ANOSIM R and P values are indicated in each figure.

GUT MICROBIOME COMPOSITION AND DAYCARE ATTENDANCE
The secondary aim of our study was to investigate the impact of daycare attendance on the community structure of the infant gut microbiome. Due to the low number of infants in our cohort we were not able to stratify each group by age. Nonetheless, we observed specific differences in EBF infants attending daycare compared to infants that were cared for at home, with significantly increased Firmicutes in the first group, specifically *Eubacterium*, Clostridiales, Ruminococcaceae, *Lactobacillus*, Roseburia, and Lachnospiraceae, and decreased Bacteroides. EBF+S infants who attended daycare had also a significant increased representation of Firmicutes, specifically *Eubacterium*, Acidaminococcus, and species of the family Clostridiaceae, as well as the Proteobacteria Sutterella. (Figure 4, Supplemental Figure 3). No non-EBF infants attended daycare; consequently we were not able to compare daycare attendance in relation to this feeding group, but we did observe an increased abundance of the taxa Faecalibacterium, Anaerotruncus, Coprobacillus, and Clostridiales in the phylum Firmicutes, Collinsella in the phylum Actinobacteria, and Enterobacteriaceae in the phylum Proteobacteria, and decreased abundance of genera of the family Ruminococcaceae in the non-EBF+S attending daycare compared to infants in the same group that did not attend daycare.

METABOLIC FUNCTIONS ASSOCIATED WITH FEEDING AND FEEDING TRANSITIONS
PICRUST analysis identified 328 different functional pathways, and showed that, despite the inter- and intra-individual variation in the composition of the fecal microbiota, samples were predicted to have similar functional profiles (Supplemental Figure 2). Overall, our analysis revealed differences in the mean relative predicted abundance of gene categories (clusters of orthologous groups, COGs) and metabolic pathways (KEGG) between feeding groups. Reflecting their lower diversity, EBF infants had the lowest number of KEGG metabolic pathways representative genes, followed by non-EBF, non-EBF+S and EBF+S infants. Samples from EBF infants, however, had an overrepresentation of pathways involved in environmental information processing, specifically membrane transport (Figure 6). EBF samples showed a significantly increased (p = 0.01) abundance of genes encoding sugar-specific phosphotransferase system (PTS) components like fructose, N acetilglucosamine, trehalose and glucose transporters. In addition the PTS system ascorbate-specific IIB component was also enriched in EBF infants.

In contrast, energy metabolism, specifically nitrogen and methane metabolism genes, were overrepresented in non-EBF and EBF+S samples compared to EBF infants (p ≤ 0.05),...
FIGURE 2 | A comparison of species richness $S$ (A) and phylogenetic diversity $PD$ (B) between infants attending daycare and infants cared for at home ($p < 0.05$). (C) PCoA analysis of samples with repeated resampling (jackknifing) according to daycare attendance regardless of feeding group. Blue symbols represent samples from infants attending daycare. Red symbols represent infants not attending daycare. ANOSIM $R$ and $P$ values are indicated in the figure.

FIGURE 3 | Relative abundance of (A) phyla and (B) genera in the different feeding groups. The mid panel shows a diagram of the taxa represented in the samples. For clarity only genera represented over 0.1% were included. The total representation per sample was $>99.9\%$. 

Thompson et al. Infant feeding and the microbiome
Table 3 | Significant differences in relative abundances (%) between feeding groups.

<table>
<thead>
<tr>
<th></th>
<th>EBF compared to non-EBF</th>
<th>EBF+S compared to non-EBF+S</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EBF</td>
<td>non-EBF</td>
</tr>
<tr>
<td>Actinobacteria*</td>
<td>48.3 ± 34.3</td>
<td>23.6 ± 28.2</td>
</tr>
<tr>
<td>Eggerthella</td>
<td>1.4e – 3</td>
<td>6.0e – 3</td>
</tr>
<tr>
<td>Bifidobacterium</td>
<td>46.4 ± 35.2</td>
<td>23.5 ± 28.2</td>
</tr>
<tr>
<td>Bacteroidetes*</td>
<td>11.0 ± 22.5</td>
<td>42.6 ± 42.7</td>
</tr>
<tr>
<td>Bacteroides</td>
<td>5.4 ± 14.6</td>
<td>42.1 ± 43.3</td>
</tr>
<tr>
<td>Firmicutes</td>
<td>37.6 ± 31.2</td>
<td>33.0 ± 25.9</td>
</tr>
<tr>
<td>Clostridiales*</td>
<td>2.7 ± 7.6</td>
<td>4.0 ± 8.1</td>
</tr>
<tr>
<td>Lachnospiraceae*</td>
<td>1.2 ± 3.3</td>
<td>8.5 ± 14.1</td>
</tr>
<tr>
<td>Blautia*</td>
<td>0.0 ± 0.0</td>
<td>0.06 ± 0.16</td>
</tr>
<tr>
<td>Faecalibacterium*</td>
<td>1.4e – 4</td>
<td>6.2e – 3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>EBF compared to EBF+S</th>
<th>non-EBF compared to non-EBF+S</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EBF</td>
<td>EBF+S</td>
</tr>
<tr>
<td>Actinobacteria*</td>
<td>48.3 ± 34.3</td>
<td>41.1 ± 34.2</td>
</tr>
<tr>
<td>Bifidobacterium</td>
<td>1.4e – 3</td>
<td>6.0e – 3</td>
</tr>
<tr>
<td>Bacteroidetes*</td>
<td>11.0 ± 22.5</td>
<td>29.0 ± 30.5</td>
</tr>
<tr>
<td>Bacteroides</td>
<td>5.4 ± 14.6</td>
<td>18.3 ± 27.1</td>
</tr>
<tr>
<td>Firmicutes</td>
<td>37.6 ± 31.2</td>
<td>29.2 ± 23.6</td>
</tr>
<tr>
<td>Peptostreptococcaceae*</td>
<td>1.4e – 3</td>
<td>6.8e – 3</td>
</tr>
<tr>
<td>Blautia*</td>
<td>0.0 ± 0.0</td>
<td>0.08 ± 0.2</td>
</tr>
<tr>
<td>Proteobacteria</td>
<td>2.93 ± 5.4</td>
<td>0.5 ± 0.74</td>
</tr>
<tr>
<td>Staphylococcus*</td>
<td>0.18 ± 0.7</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>Roseateles*</td>
<td>0.02 ± 0.04</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>Neisseria*</td>
<td>8.1e – 3</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>Escherichia/Shigella*</td>
<td>6.7e – 3</td>
<td>0.3 ± 0.7</td>
</tr>
</tbody>
</table>

Asterisks indicate p-values were ≤ 0.05; section signs indicate p ≤ 0.1. Phyla are in bold, with lower taxonomic levels indicated below.

FIGURE 4 | Relative abundance of (A) phyla and (B) genera by feeding group and attendance to daycare (D+). The mid panel shows a diagram of the taxa represented in the samples. For clarity only genera represented over 0.1% were included. The total representation per sample was >99.9%.
probably due to the higher protein content of formula and solid foods (Figure 7). Stool samples from non-EBF infants also had a significantly higher predicted abundance (p ≤ 0.05) of genes involved in protein catabolism like D-aminopeptidases, tripeptide aminopeptidases, aminocarboxylhistidine dipeptidases, dipeptidases D, and g-D-glutamyl-meso-diaminopimelate peptidases.

When we compared EBF to EBF+S samples to assess the impact of introduction of solid foods, we observed a significantly increased predicted representation of the Metabolism category, specifically biosynthesis of other secondary metabolites, energy metabolism and metabolism of terpenoids and polyketides, in EBF+S samples. Enzymatic functions significantly overrepresented in EBF+S infants included quinate/shikimate dehydrogenase (EC:1.1.1.282), involved in phenylalanine, tyrosine and tryptophan biosynthesis, oxalyl-CoA decarboxylase (EC:4.1.1.8), pyruvate ferredoxin oxidoreductase (alpha and gamma subunits) involved in several essential metabolic pathways including glycolysis, polyphosphate kinase (EC:2.7.4.1).
involved in oxidative phosphorylation and RNA degradation, starch phosphorylase (EC:2.4.1.1) involved in starch and sucrose metabolism and insulin signaling pathway, and pyruvate ferredoxin oxidoreductase (K00169 and K00172, alpha and gamma subunits), responsible for the oxidative decarboxylation of pyruvate to acetyl-coenzyme A in anaerobic bacteria (Chabriere et al., 1999).

A comparison of non-EBF and non-EBF+S samples showed an increased predicted representation of the pathways involved in cellular processes (cell growth and death), genetic information processing (replication and repair), human diseases (immune system diseases), and organismal systems (endocrine system) after the introduction of solid foods. Two hundred and thirty genes were overrepresented in non-EBF+S infants compared to non-EBF infants compared to only 24 in the EBF+S to EBF comparison group. Enzymatic functions overrepresented in the non-EBF+S included several methyl-accepting chemotaxis proteins (MCP) including serine sensor, aspartate sensor receptor, ribose and galactose sensor, and peptide sensor receptor, involved in signal transduction (Williams and Stewart, 1999), and the curli assembly protein including all components: major curlin subunit (K04334, csgA), minor curlin subunit (K04335, csgB), curli production protein (K04336, csgC), curli production assembly/transport component CsgE (K04337, csgE), curli production assembly/transport component CsgF (K04338, csgF), and curli production assembly/transport component CsgG (K06214, csgG). Curli are extracellular fibers expressed by species of *Escherichia* and *Salmonella* involved in colonization and biofilm formation (Chapman et al., 2002; Vidal et al., 1998). Also overrepresented in the non-EBF+S group compared to the non-EBF group were genes involved in the biosynthesis of colanic acid (K03207, colanic acid biosynthesis protein WcaH and K03208, colanic acid biosynthesis protein WcaI), an exopolysaccharide common to many enterobacteria (Stevenson et al., 1996) and aerobactin synthesis genes (K03894, aerobactin synthetase subunit alpha and K03895, aerobactin synthetase subunit beta), a bacterial virulence factor that allows *Escherichia coli* to sequester iron in iron-depleted environments (de Lorenzo et al., 1986).

**DISCUSSION**

The infant gut microbiota develops from birth until maturity, establishing a mutually beneficial cohabitation with the host. As a consequence, the infant gut microbiome is less stable than the adult microbiome and susceptible to modulation or disruption by environmental assault. Our study aimed to identify differences in the intestinal microbiome between exclusively breastfed and non-exclusively breastfed infants in a relatively small cohort of 9 healthy infants, who were followed from 2 weeks to 14 months of age before and after the introduction of solid foods, and between infants cared for in out-of-home daycare centers and those cared for at home. We found a lower diversity in the microbiota of...
younger infants, regardless of feeding type, which increased over time, confirming previous small and large cohorts studies that demonstrated that diversity increases from 0 to 3 years of age (Koenig et al., 2011; Yatsunenko et al., 2012). Attendance in out-of-home daycare was identified as a discriminant parameter in the composition of the gut by the ANOSIM analysis. Infants in out-of-home care showed greater diversity in the microbiota compared to those cared for in the home, regardless of feeding type. To the best of our knowledge, our study is the first to report the microbiome analysis and identification of increased species richness and diversity in infants attending daycare.

Along with these differences by age and daycare attendance, we also found significant differences across feeding patterns. Non-EBF infants displayed a higher diversity and species richness than EBF infants. Clustering analysis indicated that EBF and non-EBF groups had different overall microbiome composition. Our results confirm earlier studies (Mountzouris et al., 2002; Palmer et al., 2007; Koenig et al., 2011; Gomez-Llorente et al., 2013) showing that the gut microbiome of EBF infants was different than the microbiome of non-EBF infants, with increased proportions of *Bifidobacterium* and lower abundance of Bacteroidetes and Clostridiales. Importantly, we found that introduction of solid foods into the diet was associated with more marked changes in the microbiota composition of non-EBF infants, differences that were also reflected in the mean abundances of predicted gene categories and metabolic pathways.

In our study, non-EBF +S samples showed the highest PD values, followed by EBF +S, non-EBF and EBF samples. PCoA plots of unweighted UniFrac analysis showed that samples from EBF infants tended to segregate from samples from non-EBF infants and this tendency continued after solid foods were introduced into the diet. Interestingly, although EBF samples tended to separate from EBF +S samples, less defined clusters suggested fewer differences in the stool microbiota between these groups when solid foods were incorporated into the diet. On the contrary, we observed very defined clusters when we compared non-EBF vs. non-EBF +S samples suggesting a higher impact of solid foods on the microbiome of non-EBF babies. Many studies refer to the microbiome of formula-fed infants as “adult-like” (Harmsen et al., 2000; Olives et al., 2013); however, our data suggest that the microbiome of those infants is different to the microbiome of EBF infants as well as the adult microbiome. Further, the introduction of solid foods appeared to have a more dramatic impact in non-EBF than in EBF infants, where breast milk has prepared the host and the gut microbiome for the introduction of simple vegetal food (Matamoros et al., 2013). We hypothesize that the non-EBF microbiome could be less adaptable when faced with new food substrates, since formula composition does not vary over time while breast milk composition varies according to the mother’s diet and other factors (Coppa et al., 1993; Cabrera-Rubio et al., 2012). The apparent paradox of a low diversity, high plasticity gut microbiome certainly warrants further investigation.

We conducted Random Forests (RF) analysis to identify specific bacterial taxa highly predictive of feeding-type and feeding transition. RF has been recently used in microbiome analysis because RF classification models are versatile, have a high prediction accuracy and provide additional information such as variable importance (Touw et al., 2013). RF identified 60 highly predictive bacterial taxa with 28% from the the Actinobacteria, 12 specifically from the genus *Bifidobacterium*. The dominant taxa in EBF infants were Actinobacteria followed by Firmicutes whereas the microbiome of non-EBF infants was dominated by Bacteroidetes and Firmicutes. Specifically, EBF infants had a higher relative abundance of *Bifidobacterium* species compared to non-EBF infants where *Bacteroides* species were dominant. Dominance of *Bifidobacterium* in EBF vs. non-EBF infants has been debated since numerous studies have found a lower abundance of *Bifidobacterium* in formula-fed infants relative to breast-fed infants (Yoshioka et al., 1983; Favier et al., 2002; Hopkins et al., 2005), but other studies have reported *Bifidobacterium* as the dominant group in both groups (Yatsunenko et al., 2012; Gomez-Llorente et al., 2013). Considering the high interindividual variation we might hypothesize that both groups are Colonized early by *Bifidobacterium*, but the abundance of these microorganisms is not favored by formula, while it is maintained in exclusively breastfed babies by the presence of oligosaccharides in human milk.

Discernible predicted metabolic differences were observed between EBF and non-EBF infants, with increased representation of sugar transporters in EBF samples compared to non-EBF samples, which are essential for the digestion of milk carbohydrates (Candela et al., 2012). Conversely, non-EBF samples were enriched in genes involved in energy metabolism, specifically nitrogen and methane metabolism genes as well as peptidases, probably due to the higher protein content of formula. The PTS system ascorbate-specific IIB component was also enriched in EBF infants. This is in accordance with different studies that showed a higher content of vitamin C in breast fed infants than formula fed infants (Tlaskal and Novakova, 1990; Aycicek et al., 2006). Kurokawa et al. (2007) observed in an early study that the infant microbiome was dominated by COGs included in the Carbohydrate transport and metabolism, including 12 families of glycosylhydrolases. They also identified several enzymes involved in degradation of non-digestible polysaccharides of plant origin.

Human milk is composed of lactose (70 g/L), followed by lipids (40 g/L). Interestingly, the third most abundant component are human milk oligosaccharides (HMO; 5–15 g/L), followed by protein at 8 g/L (Zivkovic et al., 2011). When infants switch from breastfeeding to solids, the fraction of energy from protein increases rapidly (Zivkovic et al., 2011). Accordingly, when we compared EBF to EBF +S samples to assess the impact of introduction of solid foods, we observed a significantly increased representation of the Metabolism category, specifically biosynthesis of other secondary metabolites, energy metabolism and metabolism of terpenoids and polyketides in EBF +S samples. Yet, as reflected in our UniFrac PCoA analysis that suggested subdued changes in the microbiome of EBF infants when solids were introduced into their diet, only 24 enzymes were over represented in EBF +S infants compared to the EBF group. Of interest are the overrepresentation of the biotin operon repressor BirA (K03524), which regulates transcription of the prokaryotic biotin operon (Wilson et al., 1992), the thiamine biosynthesis lipoprotein (K03734) involved in the conversion of aminomimidazole ribotide.
(AIR), a purine intermediate, to the 4-amino-5-hydroxymethyl-2-methyl pyrimidine (HMP) moiety of thiamine in Salmonella (Beck and Downs, 1998), and the oxalyl-CoA decarboxylase involved in oxalate degradation (Azcarate-Peril et al., 2006). Overrepresentation of these pathways reflects the adaptation of the microbiome to availability of micronutrients due to introduction of new foods.

In contrast, over 200 bacterial gene categories were overrepresented in non-EBF+S compared to non-EBF infants including several methyl-accepting chemotaxis proteins (MCP) including serine sensor, aspartate sensor receptor, ribose and galactose sensor, and peptide sensor receptor, involved in signal transduction (Williams and Stewart, 1999). Signal-responsive components of transmembrane signal-transducing regulatory systems include methyl-accepting chemotaxis proteins and membrane-bound, two-component histidine kinases. Prokaryotes use these regulatory networks to channel environmental cues into adaptive responses. Also overrepresented were all curlin components, extracellular fibers expressed by species of Escherichia and Salmonella involved in colonization and biofilm formation (Vidal et al., 1998; Chapman et al., 2002). Such pathway differences may underlie the long standing findings that breastfeeding is associated with lower morbidity from a number of diseases including acute otitis media, atopic dermatitis, gastrointestinal infections, lower respiratory tract diseases, asthma, obesity, cardiovascular diseases, type 1 and type 2 diabetes, childhood leukemia, and Sudden Infant Death Syndrome (SIDS) in infants in developed countries (Ip et al., 2007).

Despite the limitations of PICRUSt as a predictor of gene functionality from 16S amplicon sequencing data, we found similarities between our results and previously reported metagenomic sequencing (Koenig et al., 2011; Vaishampayan et al., 2011; Schwartz et al., 2012). Two previous studies reported the overrepresentation of genes involved in sugar transport of human milk glycans like phosphotransferase system for N-acetylglucosamine and glucose and for other carbohydrates like fructose in EBF infants (Koenig et al., 2011; Vaishampayan et al., 2011). Koenig et al. (2011) also reported the enrichment of genes involved in vitamin and cofactor biosynthesis after the introduction of solids in one EBF infant. This is in accordance with our results in which we detected the overrepresentation of genes involved in biosynthesis of biotine and thiamine when solids were introduced into the diet of EBF infants. Finally, Schwartz et al. (2012), found virulence characteristics that differed between breast-fed and formula fed infants. Among others, they detected an enrichment of genes involved in iron scavenging. This is in agreement with our report of an over-representation of genes involved in the synthesis of aerobactin, an iron-chelating agent, in non-EBF+s solids infants.

Our collection of infant stool samples as part of prospective, longitudinal study of “normal” infant growth and development expands previous descriptive research (Koenig et al., 2011). Further, the high rates of breastfeeding permitted examination of the effects of mixed breast- and formula-feeding, a group that has been less often described in the literature, and the impact of solid food introduction in EBF and non-EBF infants. Our use of 24-h food diaries at each sample collection likely reduced misclassification of infant feeding type based on recall. Along with these strengths, the present study has a number of limitations. Due to the small sample available for secondary analysis, we were not able to test the impact of other characteristics known to shape the infant microbiome, such as birth type (vaginal vs. Caesarian-section) or antibiotic use. While this sample size is relatively small, our fine-grained analysis permits hypothesis generation for larger studies to assess the bacterial and functional impacts of solid food introduction and daycare attendance. Additionally, we used PICRUSt to estimate metagenomic information rather than directly measuring the metagenome. Consequently, our findings of metabolic differences between the groups reflect predicted, rather than actual, function; nevertheless, PICRUSt results have been validated against metagenomic sequencing and show correlations of 80–90% in human studies (Langille et al., 2013).

Our study showed that feeding and feeding-transitions have significant, dramatic impacts on the infant gut microbiome and that out-of-home day care attendance may further influence microbial diversity independently of feeding. These results indicate that the early nutrient and pathogenic environment may shape colonization during this highly plastic and sensitive period. Our findings that different genetic and metabolic pathways were activated in EBF and non-EBF infants, even among infants that were still predominantly receiving breastmilk, further suggest that feeding-based differences in microbiome composition have the potential to contribute to the programming of infant metabolism and immune function. Such programming during the critical period of weaning with the transition from a milk-based to a more varied, solid-based diets characterized by higher levels of carbohydrates and animal proteins may have long-term consequences for not only the establishment of the adult microbiome but also the development of metabolic diseases like obesity, diabetes, cardiovascular disease, and non-alcoholic fatty liver disease (NAFLD).

AUTHOR CONTRIBUTIONS
All authors contributed to study design and interpretation. Amanda L. Thompson and Michelle L. Lampl acquired the samples. Andrea Monteagudo-Mera and Maria B. Cadenas conducted DNA isolation and sample processing for sequencing. Andrea Monteagudo-Mera and M. A. Azcarate-Peril conducted data analysis. Amanda L. Thompson, Andrea Monteagudo-Mera and M. A. Azcarate-Peril drafted the manuscript. All authors provided critical revision to the manuscript and have approved the final version. All authors agree to take responsibility for accuracy and integrity of the research.

ACKNOWLEDGMENTS
We are thankful to Allison McFadden for insightful scientific discussions. The Microbiome Core Facility is supported in part by the NIH/National Institute of Diabetes and Digestive and Kidney Diseases grant P30 DK34987. Funding for this analysis was provided in part by the University Research Council, University of North Carolina at Chapel Hill. Amanda L. Thompson is supported by NIH NICHD grant K01 HD071948-01. We thank the Carolina Population Center at the University of North Carolina at Chapel Hill (NIH 5 R24 HD050924) for general support.
SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fcimb.2015.00033/abstract

Supplemental Figure 1 | Bacterial Phylogenetic Diversity (PD) and species richness (S) during the first 14 months of age increases over time.

Supplemental Figure 2 | Relative abundance of bacterial phyla (bacterial phylogenetic diversity, PD), bacterial gene functional representation, feeding type and collection time points in stool samples by individual infant.

Supplemental Figure 3 | Relative abundances of bacterial taxa over- and under-represented in infants attending daycare by feeding group.

Supplemental Table 1 | List of predictive OTUs generated by Random forests–supervised classification analysis.

REFERENCES


Tannock, G. W., Lawley, B., Munro, K., Gowri Pathmanathan, S., Zhou, S. J., Makrides, M., et al. (2013). Comparison of the compositions of the stool micro-


**Conflict of Interest Statement**: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.